



PCT

RAW SEQUENCE LISTING

DATE: 04/03/2003

PATENT APPLICATION: US/09/719,533A

TIME: 14:24:29

Input Set : A:\U013108-9sequence.txt

Output Set: N:\CRF4\04032003\I719533A.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Chen, Chong Jin
 6 Lin, Gex Heow
 7 Leong, Ai Lin
 8 Thiao, Yi
 9 Chen, Wei Ning

11 (ii) TITLE OF INVENTION: A VACCINE-INDUCED HEPATITIS B VIRAL
 12 CTFAIN AND USES THEREOF

14 (iii) NUMBER OF SEQUENCES: 11

16 (iv) CORRESPONDENCE ADDRESS:

17 (A) ADDRESSEE: Ladani & Parry
 18 (B) STREET: 16 West 61 Street
 19 (C) CITY: New York
 20 (D) STATE: New York
 21 (E) COUNTRY: USA
 22 (F) ZIP: 10003

24 (v) COMPUTER READABLE FORM:

25 (A) MEDIUM TYPE: Floppy disk
 26 (B) COMPUTER: IBM PC compatible
 27 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 28 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

30 (vi) CURRENT APPLICATION DATA:

C--> 31 (A) APPLICATION NUMBER: US/09/719,533A
 C--> 32 (B) FILING DATE: 13-Dec-2000
 33 (C) CLASSIFICATION: 485

35 (vii) PRIOR APPLICATION DATA:

36 (A) APPLICATION NUMBER: PCT/SG98/00045
 37 (B) FILING DATE: 19-JUN-1998

39 (viii) ATTORNEY/AGENCY INFORMATION:

40 (A) NAME: Miller, Cynthia E.
 41 (B) REGISTRATION NUMBER: 34,678
 42 (C) REFERENCE/DOCKET NUMBER: U-013108-9

44 (ix) TELECOMMUNICATION INFORMATION:

45 (A) TELEPHONE: 212-703-1890

47 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 215 base pairs
 51 (B) TYPE: nucleic acid
 52 (C) STRANDEDNESS: double
 53 (D) TOPOLOGY: circular

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

57 CTCCACCACT TTCCAGTAAA CTCTTCAAGA TCCAGAGTC AGGGCCCTGT ACTTTCCTGC

60

ENTERED

RAW SEQUENCE LISTING

DATE: 04/02/2003

PATENT APPLICATION: US/09/719,533A

TIME: 14:14:29

Input Set : A:\U013108-9sequence.txt

Output Set: N:\CRF4\04032003\I719533A.raw

```

59 TGGTGGCTTC AGTTCAGGAA CAGTGAGGAC TGCTCAGAAT ACTGTCTCTG CCATATCTTC 120
61 AATCTTATCG AAGACTGGGG ACCCTGTACG GAACATGGAG AACATCGCAT CAGGACTCTT 180
63 AGGACCCCTG CTCTGTTCAT AGGCGGGGTT TTCTTGTTCG AAAAAATCT TCACAATAGC 240
65 GCAGAGTCTA GATCTGTGCT GGACTTCCTT CAATTTCTTA GGCGGAGACG CCGTCTCTCT 300
67 TGGGCAAAAT TCGAGTCTCT AATCTCTAG TCACTCAGCA ACCTGTTCTC CCGCATCTTC 360
69 TCGTGGTTAT CGCTGGATCT GTCTGCTGGG TCTTATCTTC TCTCTCTGCA TCGTCTCT 420
71 ATGTTCATCG TCTTCTTCG TCTCTCTGCA CTATCAAGGT ATCTTGGCG TCTCTCTCT 480
73 AATTTCAGGA TCACTACAAA CAGGACGG AGCATGAAAA ACCTGAGACA CTCTCTCT 540
75 AGGAACTCTT AATCTCTCT CATCTCTCTG TACAAAACT ACCTGAGACA CTCTCTCT 600
77 TATCTCTCTC CACTATCTCT GGGCTTCTG AAAAACTA TGAGAGTG CTCTCTCT 660
79 TTTCTCTCTG CTCTCTCTT TATCTCTCT TCTCTCTG TCTCTCTG TCTCTCTCT 720
81 TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 780
83 GAGTCTCTCT ATCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 840
85 CAGAAAACTA AAACTCTCTCT ATCTCTCTCT AACTCTCTCT GACTCTCTCT TCTCTCTCT 900
87 GCTCTCTCT CACTCTCTCT TATCTCTCT AAAAACTA TCTCTCTCT TCTCTCTCT 960
89 GAAAACTCT CACTCTCTCT GAAAACTCT CAACTCTCT TCTCTCTCT TCTCTCTCT 1020
91 GCTCTCTCT CACTCTCTCT ATCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1080
93 AAACTCTCT TCTCTCTCT CACTCTCTCT AACTCTCTCT TCTCTCTCT TCTCTCTCT 1140
95 CTTCTCTCT TCTCTCTCT AACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1200
97 ACTCTCTCT GCTCTCTCT AACTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT 1260
99 CCACTCTCT CACTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1320
101 CTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1380
103 CTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1440
105 CTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1500
107 CTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1560
109 CTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1620
111 TCACTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1680
113 TCTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1740
115 TCTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1800
117 CTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1860
119 CTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1920
121 AGAATCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 1980
123 TCTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 2040
125 TCTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 2100
127 TCTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 2160
129 TCTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 2220
131 TCTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 2280
133 TCTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 2340
135 TCTCTCTCT CACTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT TCTCTCTCT 2400

```

RAW SEQUENCE LISTING

DATE: 04/03/2002

PATENT APPLICATION: US/09/719,533A

TIME: 14:24:29

Input Set : A:\U013108-9sequence.txt

Output Set: N:\CRF4\04032003\I719533A.raw

```

157 TGGCCGGACG CCAACAAGGT GGGASTGGGA GCATTGGGGC CAGGGTTTAC CCTCTTCAT 3060
159 GGGGGACTGT TGGGGTGGAG CCTTCAGCT CAGGGCTAC TCACAACTGT GCCAGCAGCT 3120
161 CCTCTCTCTG CCTCCACCAA TCAGCACTCA GGAAGGCAGC CTACTCCTT ATCTCCAGCT 3180
163 CTAAGGGACA CTCATCTCA GGCATTCAG TGGAA 3215
164 (2) INFORMATION FOR SEQ ID NO: 2:
165 (i) SEQUENCE CHARACTERISTICS:
166 (A) LENGTH: 843 amino acids
167 (B) TYPE: amino acid
168 (C) STRANDEDNESS: single
169 (D) TOPOLOGY: linear
170 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
171 Met Pro Leu Ser Tyr Gln His Phe Arg Lys Leu Leu Leu Leu Asp Glu
172 1 5 10 15
173 Glu Ala Gly Pro Leu Glu Glu Glu Leu Pro Arg Leu Ala Asp Glu Gly
174 20 25 30
175 Leu Asn Arg Arg Val Ala Glu Asp Leu Asn Leu Gly Asn Leu Asn Val
176 35 40 45
177 Ser Ile Pro Trp Thr His Lys Val Gly Asn Phe Thr Gly Leu Tyr Ser
178 50 55 60
179 Ser Thr Val Pro Cys Phe Asn Pro Lys Trp Gln Thr Pro Ser Phe Pro
180 65 70 75 80
181 Asp Ile His Leu Gln Glu Asp Ile Leu Asp Arg Cys Lys Gln Phe Val
182 85 90 95
183 Glu Pro Leu Thr Val Asn Glu Asn Arg Arg Leu Lys Leu Ile Met Pro
184 100 105 110
185 Ala Arg Phe Tyr Pro Asn Val Thr Lys Tyr Leu Pro Leu Asp Lys Gly
186 115 120 125
187 Ile Lys Pro Tyr Tyr Pro Glu Tyr Val Val Asn His Tyr Pro Gln Thr
188 130 135 140
189 Arg His Tyr Leu His Thr Leu Trp Lys Ala Gly Ile Leu Tyr Lys Arg
190 145 150 155 160
191 Gln Ser Thr Arg Ser Ala Ser Phe Cys Gly Ser Pro Tyr Ser Trp Glu
192 165 170 175
193 Gln Asp Leu Gln His Gly Arg Leu Val Phe Gln Thr Ser Lys Arg His
194 180 185 190
195 Gly Asp Lys Ser Phe Cys Pro Glu Ser Pro Gly Ile Leu Pro Arg Ser
196 195 200 205
197 Ser Val Gly Pro Cys Ile Gln Ser Gln Leu Arg Lys Ser Arg Leu Gly
198 210 215 220
199 Pro Gln Pro Ala Gln Gly Gln Leu Ala Gly Arg Gln Gln Gly Gly Ser
200 225 230 235 240
201 Gly Ser Ile Arg Ala Arg Val His Pro Ser Ser Trp Gly Thr Val Gly
202 245 250 255
203 Val Glu Pro Ser Gly Ser Gly Pro Thr His Asn Cys Ala Ser Ser Ser
204 260 265 270
205 Ser Ser Cys Leu His Gln Ser Ala Val Arg Lys Ala Ala Tyr Ser Leu
206 275 280 285
207 Ile Ser Thr Ser Lys Gly His Ser Ser Ser Gly His Ala Val Glu Leu
208 290 295 300

```

RAW SEQUENCE LISTING

DATE: 04/03/2002

PATENT APPLICATION: US/09/719,533A

TIME: 14:24:29

Input Set : A:\U013108-9sequence.txt

Output Set: N:\CRF4\04032003\I719533A.raw

225	His	His	Phe	Pro	Pro	Asn	Ser	Ser	Arg	Ser	Gln	Ser	Gln	Gly	Pro	Val
226	305					210				315						320
228	Leu	Ser	Cys	Trp	Trp	Leu	Gln	Phe	Arg	Asn	Ser	Glu	Pro	Cys	Ser	Glu
229				325						330					335	
241	Tyr	Cys	Leu	Cys	His	Ile	Val	Asn	Leu	Ile	Glu	Asp	Trp	Gly	Pro	Cys
242				340					345					350		
244	Thr	Glu	His	Gly	Glu	His	Arg	Ile	Arg	Thr	Pro	Arg	Thr	Pro	Ala	Arg
245			355					360					365			
247	Val	Thr	Gly	Gly	Val	Phe	Leu	Val	Asp	Lys	Asn	Pro	His	Asn	Thr	Ala
248			370				375					380				
250	Glu	Ser	Arg	Leu	Val	Val	Asp	Phe	Ser	Gln	Phe	Ser	Arg	Gly	Asn	Thr
251	385					390				395						400
253	Arg	Val	Ser	Trp	Pro	Lys	Phe	Ala	Val	Pro	Asn	Leu	Gln	Ser	Leu	Thr
254				405						410					415	
256	Asn	Leu	Leu	Ser	Ser	Asn	Leu	Ser	Trp	Leu	Ser	Leu	Asp	Val	Ser	Ala
257				420					425					430		
259	Ala	Phe	Tyr	His	Leu	Pro	Leu	His	Pro	Ala	Ala	Met	Pro	His	Leu	Leu
260			435					440					445			
262	Val	Gly	Ser	Ser	Gly	Leu	Ser	Arg	Tyr	Val	Ala	Arg	Leu	Ser	Ser	Asn
263			450				455					460				
265	Ser	Arg	Ile	Asn	Asn	Asn	Glu	His	Arg	Thr	Met	Glu	Asn	Leu	His	Asn
266				465		470				475					480	
268	Ser	Cys	Ser	Asp	Asp	Leu	Tyr	Val	Ser	Leu	Met	Leu	Leu	Tyr	Lys	Thr
269				485						490					495	
271	Tyr	Gly	Gln	Lys	Leu	His	Leu	Tyr	Ser	His	Pro	Ile	Ile	Leu	Gly	Phe
272				500					505					510		
274	Arg	Lys	Ile	Pro	Met	Gly	Val	Gly	Leu	Ser	Pro	Phe	Leu	Leu	Ala	Gln
275			515					520					525			
277	Phe	Thr	Ser	Ala	Ile	Cys	Ser	Val	Val	Arg	Arg	Ala	Phe	Pro	His	Cys
278			530				535					540				
280	Leu	Ala	Phe	Ser	Tyr	Met	Asp	Asp	Val	Val	Leu	Gly	Ala	Lys	Ser	Val
281				545			550				555				560	
283	Gln	His	Leu	Gln	Ser	Leu	Tyr	Ala	Ala	Val	Thr	Asn	Phe	Leu	Leu	Ser
284				565						570					575	
286	Leu	Gly	Ile	His	Leu	Asn	Pro	His	Lys	Thr	Lys	Arg	Trp	Gly	Tyr	Ser
287				580					585					590		
289	Leu	Asn	Phe	Met	Gly	Tyr	Val	Ile	Gly	Ser	Trp	Gly	Thr	Leu	Pro	Gln
290			595					600					605			
292	Glu	His	Ile	Val	Gln	Lys	Ile	Lys	Met	Cys	Phe	Arg	Lys	Leu	Pro	Val
293			610				615					620				
295	Asn	Arg	Pro	Ile	Asp	Trp	Lys	Val	Cys	Gln	Arg	Ile	Val	Gly	Leu	Leu
296			625				630				635				640	
298	Gly	Phe	Ala	Ala	Pro	Phe	Thr	Gln	Cys	Gly	Tyr	Pro	Ala	Leu	Met	Pro
299				645						650					655	
301	Leu	Tyr	Ala	Cys	Ile	Gln	Ala	Lys	Gln	Ala	Phe	Thr	Phe	Ser	Gln	Thr
302				660					665					670		
304	Tyr	Lys	Thr	Phe	Leu	Ser	Lys	Gln	Tyr	Leu	Asn	Leu	Tyr	Pro	Val	Ala
305				675				680					685			
308	Arg	Gln	Arg	Pro	Gly	Leu	Cys	Glu	Val	Phe	Ala	Asp	Ala	Thr	Pro	Thr

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/719,533A

DATE: 04/03/2003

TIME: 14:24:29

Input Set : A:\U013108-9sequence.txt

Output Set: N:\CRF4\04032003\I719533A.raw

```

309          690          695          700
311 Gly Trp Gly Leu Ala Ile Gly His Gln Arg Met Arg Gly Thr Phe Val
312 705          710          715          720
314 Ser Pro Leu Pro Ile His Thr Ala Glu Leu Leu Ala Ala Cys Phe Ala
315          725          730          735
317 Arg Ser Arg Ser Gly Ala Lys Leu Ile Gly Thr Asp Asn Ser Val Val
318          740          745          750
320 Leu Ser Arg Lys Tyr Thr Ser Phe Pro Trp Leu Leu Gly Cys Ala Ala
321          755          760          765
323 Asn Trp Ile Leu Arg Gly Thr Ser Phe Val Tyr Val Pro Ser Ala Leu
324          770          775          780
326 Asn Pro Ala Asp Asp Pro Ser Arg Gly Arg Leu Gly Leu Tyr Arg Pro
327          785          790          795          300
329 Leu Leu Arg Leu Leu Tyr Arg Pro Thr Thr Gly Arg Thr Ser Leu Tyr
330          800          810          815
332 Ala Asp Ser Pro Ser Val Pro Ser His Leu Pro Asp Arg Val His Phe
333          820          825          830
335 Ala Ser Pro Leu His Val Ala Trp Arg Pro Pro
336          835          840
340 (2) INFORMATION FOR SEQ ID NO: 3:
341 (i) SEQUENCE CHARACTERISTICS:
342 (A) LENGTH: 460 amino acids
343 (B) TYPE: amino acid
344 (C) STRANDEDNESS: single
345 (D) TOPOLOGY: linear
346 (ii) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
347 Met Gly Gly Trp Ser Ser Lys Pro Arg Lys Gly Met Gly Thr Asn Leu
348 1 5 10 15
349 Ser Val Pro Asn Pro Leu Gly Phe Phe Pro Asp His Gln Leu Asp Pro
350 20 25 30
351 Ala Phe Lys Ala Asn Ser Gln Asn Pro Asp Trp Asp Leu Asn Pro His
352 35 40 45
353 Lys Asp Asn Trp Pro Asp Ala Asn Lys Val Gly Val Gly Ala Phe Gly
354 50 55 60
355 Pro Gly Phe Thr Pro Pro His Gly Gly Leu Leu Gly Trp Ser Pro Gln
356 65 70 75 80
357 Ala Gln Gly Leu Leu Thr Thr Val Pro Ala Ala Pro Pro Pro Ala Ser
358 85 90 95
359 Thr Asn Arg Gln Ser Gly Arg Gln Pro Thr Pro Leu Ser Pro Pro Leu
360 100 105 110
361 Arg Asp Thr His Pro Gln Ala Met Gln Trp Asn Ser Thr Thr Phe His
362 115 120 125
363 Gln Thr Leu Gln Asp Pro Arg Val Arg Ala Leu Tyr Phe Pro Ala Gly
364 130 135 140
365 Gly Ser Ser Ser Gly Thr Val Ser Pro Ala Gln Asn Thr Val Ser Ala
366 145 150 155 160
367 Ile Ser Ser Ile Leu Ser Lys Thr Gly Asp Pro Val Pro Asn Met Glu
368 165 170 175
369 Asn Ile Ala Ser Gly Leu Leu Gly Pro Leu Leu Val Leu Gln Ala Gly

```

VERIFICATION SUMMARY

DATE: 04/03/2003

PATENT APPLICATION: US/09/719,533A

TIME: 14:24:30

Input Set : A:\U013108-9sequence.txt

Output Set: N:\CRF4\04032003\I719533A.raw

L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]